

NUCLEOTIDE SEQUENCES CODING FOR POLYPEPTIDES ENDOWED WITH A  
LARVICIDAL ACTIVITY TOWARDS LEPIDOPTERA

---

5                   The subject of the invention is nucleotide sequences coding  
for polypeptides endowed with a larvicidal activity towards  
Lepidoptera.

10                   It relates more particularly to agents, in particular  
nucleotide sequences, polypeptides or even vectors, or bacterial  
strains modified by these sequences and expressing polypeptides making  
it possible to prepare larvicidal compositions active against  
Lepidoptera, preferably against Spodoptera littoralis (hereafter  
S.littoralis) or Mamestra brassicae (hereafter designated by  
M.brassicae) or capable of transforming the plants to be treated in  
15                   conferring on them this type of activity.

                  It is known that most of the isolates of B.thuringiensis  
show a toxic activity with regard to larvae of more than a hundred  
species of Lepidoptera.

20                   This activity results from the capacity of the strains of  
B.thuringiensis to synthesize, at the moment of sporulation,  
crystalline inclusions of protein nature, or  $\delta$ -endotoxins, under the  
control of one or several types of gene.

                  It has been shown that the activity of these polypeptides  
is contained in the  $\text{NH}_2$ -terminal half or N-terminus of the protein.

25                   The studies carried out have shown the high specificity  
of the  $\delta$ -endotoxins towards larvae of a given species.

                  On account of this high specificity, many species of  
Lepidoptera, in particular of the family of the Noctuidae, react only  
weakly to commercial preparations of available B.thuringiensis.

30                   It is so in particular for the species S.littoralis, a poly-  
phagous insect which constitutes the principal parasite of cotton  
and other industrially important crops. Among these crops, mention  
should be made of maize, the castor oil plant, tobacco, the groundnut,  
fodder plants, such as clover or alfalfa, or also market garden produce  
35                   such as the cabbage or the tomato.

filed  
12/11/89

Hence, one can imagine the interest of disposing of agents targeting specifically and effectively the family of the Noctuidae and in particular S.littoralis or M.brassicae.

The genes for  $\delta$ -endotoxins hitherto identified do not code for a polypeptide preferentially active with regard to S.littoralis.

The search by the inventors for a sequence of nucleotides coding for a polypeptide preferably active against the Noctuidae, more especially against S.littoralis, has led them to study the natural isolates of two strains of B.thuringiensis, the larvicidal activity of which on S.littoralis appears to be higher than that of the industrial preparations made starting from other strains of B.thuringiensis.

The species in question are aizawai 7-29 and entomocidus 6-01.

The study of these isolates has made it possible to demonstrate the existence of several genes for  $\delta$ -endotoxins of different structures and different specificities, of which two genes preferentially active against P.brassicae but not very active against the Noctuida of cotton and a gene inactive against P.brassicae and S.littoralis.

By studying the total DNA of these isolates and by carrying out appropriate hybridizations, followed by the cloning of the fragments identified by hybridization, the inventors have observed that it is possible to isolate nucleotide sequences implicated in genes for  $\delta$ -endotoxins coding for polypeptides active, preferably, against S.littoralis.

Thus, the aim of the invention is to provide nucleotide sequences capable of coding for at least the  $\text{NH}_2$ -terminal part of a  $\delta$ -endotoxin toxic against the Noctuidae and preferably against S.littoralis or M.brassicae.

It also has the aim of providing a polypeptide toxic with regard to the Noctuidae.

Furthermore, the invention relates to a procedure for obtaining such a sequence and a polypeptide showing the desired activity as well as the intermediate agents such as vectors and

bacterial strains which can be utilized for obtaining the polypeptide.

In addition, the invention relates to the uses of these sequences and polypeptides for the development of larvicidal compositions with regard to the Noctuidae, in particular S.littoralis and for the transformation of the plants likely to be infected by these larvae.

The invention relates to a sequence of nucleotides coding for at least a part of the N-terminal region of a polypeptide toxic specifically against the larvae of Lepidoptera of the Noctuidae family, and preferably against S.littoralis, characterized by its capacity of hybridization with a gene capable of expressing a polypeptide toxic towards larvae of S.littoralis.

According to another aspect of the invention, the nucleotide sequence is characterized in that it is carried by a sequence of nucleotides of about 3 kb such as obtained by in vitro genetic recombination of sequences of nucleotides of B.thuringiensis capable of hybridizing with probes 1, 2 and 3 of pHTA2 shown in figure 2. The fragment of 3 kb corresponds more particularly to the restriction fragment HindIII-PstI.

The sequences of nucleotides of the invention are, in addition, characterized in that they contain sites in the following order : HindIII - HincII - BglIII - KpnI - HindIII - PstI.

In a preferred manner, these sequences of nucleotides are obtained by in vitro genetic recombination of DNA sequences derived from at least one strain of B.thuringiensis. In a variant of the embodiment of the invention, two different strains of B.thuringiensis are utilized.

Strains of B.thuringiensis particularly suited for obtaining these sequences of nucleotides are the strains corresponding to aizawai 7-29 and entomocidus 6-01, deposited on 21 April 1987 under the No. I-661 and No. I-660, respectively, with the National Collection of Cultures of Microorganisms (N.C.C.M.) in Paris.

In an advantageous manner, the sequences of nucleotides of the invention code for a polypeptide capable of forming an immunological complex with antibodies directed against polypeptides showing the larvicidal activity with regard to S.littoralis.

A sequence of nucleotides according to the invention is characterized in that it has the capacity to hybridize with a probe formed from the sequence (I) showing the following chain arrangement:

```

      52
GTC TAC TTG ACA GGG GTA GCA ACA TAA TCG GTC AAT TTT AAA
      112
TAT GGG GCA TAT ATT GAT ATT TTA TAA AAT TTG TTA CGT TTT
      172
TTG TAT TTT TTC ATA ACA TGT GTC ATA TGT ATT AAA TCG TGG
TAA TGA AAA ACA GTA TCA AAC TAT CAG AAC TTT GGT AGT TTA
      232
ATA AAA AAA CGG AGG TAT TTT ATG GAG GAA AAT AAT CAA AAT
      292
CAA TGC ATA CCT TAC AAT TGT TTA AGT AAT CCT GAA GAA GTA
CTT TTG GAT GGA GAA CGG ATA TCA ACT GGT AAT TCA TCA ATT
      352
GAT ATT TCT CTG TCA CTT GTT CAG TTT CTG GTA TCT AAC TTT
      412
GTA CCA GGG GGA GGA TTT TTA GTT GGA TTA ATA GAT TTT GTA
TGG GGA ATA GTT GGC CCT TCT CAA TGG GAT GCA TTT CTA GTA
      472
CAA ATT GAA CAA TTA ATT AAT GAA AGA ATA GCT GAA TTT GCT
      532
AGG AAT GCT GCT ATT GCT AAT TTA GAA GGA TTA GGA AAC AAT
      592
TTC AAT ATA TAT GTG GAA GCA TTT AAA GAA TGG GAA GAA GAT
CCT AAT AAT CCA GAA ACC AGG ACC AGA GTA ATT GAT CGC TTT
      652
CGT ATA CTT GAT GGG CTA CTT GAA AGG GAC ATT CCT TCG TTT
      712
CGA ATT TCT GGA TTT GAA GTA CCC CTT TTA TCC GTT TAT GCT
CAA GCG GCC AAT CTG CAT CTA GCT ATA TTA AGA GAT TCT GTA
      772
ATT TTT GGA GAA AGA TGG GGA TTG ACA ACG ATA AAT GTC AAT
      832
GAA AAC TAT AAT AGA CTA ATT AGG CAT ATT GAT GAA TAT GCT
      892
GAT CAC TGT GCA AAT ACG TAT AAT CGG GGA TTA AAT AAT TTA
CCG AAA TCT ACG TAT CAA GAT TGG ATA ACA TAT AAT CGA TTA
      952
CGG AGA GAC TTA ACA TTG ACT GTA TTA GAT ATC GCC GCT TTC
TTT CCA AAC TAT GAC

```

Sequences of nucleotides coding for at least a part of the N-terminal region of a polypeptide toxic specifically towards larvae of Lepidoptera of the Noctuidae family, and preferably towards S.littoralis, are characterized in that they contain the chain arrangement (I) defined above.

In an advantageous manner, the sequence of nucleotides characterized by the chain arrangement defined above codes for a part of a polypeptide having a higher larvicidal activity towards S.littoralis than that of the polypeptides encoded by natural isolates presently known for their effects against S.littoralis.

The study of this sequence of nucleotides shows that it is characterized by an initiation codon ATG situated at position 241 starting from which an open reading frame of 750 nucleotides has been identified.

This sequence is also characterized by a GGAGG attachment site for ribosomes at positions 230 to 234.

According to another feature, the sequence of nucleotides of the invention is characterized in that it contains, upstream from the ATG codon, a sequence going from the nucleotide at position 137 to the nucleotide at position 177, strongly homologous with the region found by Wong et al. (1983) and described in (16) upstream from the gene for the crystal of the strain kurstaki HD1 Dipel (BTK) and for which the authors have shown that it contains three promoters BtI, BtII and Ec which are functional in B.thuringiensis and E.coli, respectively. The homology of these sequences is about 70%.

The invention also relates to a sequence of nucleotides coding for the following sequence (II) of amino acids :

MET GLU GLU ASN ASN GLN ASN  
GLN CYS ILE PRO TYR ASN CYS LEU SER ASN PRO GLU GLU VAL  
LEU LEU ASP GLY GLU ARG ILE SER THR GLY ASN SER SER ILE  
ASP ILE SER LEU SER LEU VAL GLN PHE LEU VAL SER ASN PHE  
VAL PRO GLY GLY PHE LEU VAL GLY LEU ILE ASP PHE VAL TRP  
GLY ILE VAL GLY PRO SER GLN TRP ASP ALA PHE LEU VAL GLN  
ILE GLU GLN LEU ILE ASN GLU ARG ILE ALA GLU PHE ALA ARG  
ASN ALA ALA ILE ALA ASN LEU GLU GLY LEU GLY ASN ASN PHE  
ASN ILE TYR VAL GLU ALA PHE LYS GLU TRP GLU GLU ASP PRO  
ASN ASN PRO GLU THR ARG THR ARG VAL ILE ASP PRG PHE ARG  
ILE LEU ASP GLY LEU LEU GLU ARG ASP ILE PRO SER PHE ARG  
ILE SER GLY PHE GLU VAL PRO LEU LEU SER VAL TYR ALA GLN  
ALA ALA ASN LEU HIS LEU ALA ILE LEU ARG ASP SER VAL ILE  
PHE GLY GLU ARG TRP GLY LEU THR THR ILE ASN VAL ASN GLU  
ASN TYR ASN ARG LEU ILE ARG HIS ILE ASP GLU TYR ALA ASP  
HIS CYS ALA ASN THR TYR ASN ARG GLY LEU ASN ASN LEU PRO  
LYS SER THR TYR GLN ASP TRP ILE THR TYR ASN ARG LEU ARG  
ARG ASP LEU THR LEU THR VAL LEU ASP ILE ALA ALA PHE PHE  
PRO ASN TYR ASP

A better identification of the sequence of nucleotides isolated from the above strains, deposited with the NCCM has made it possible to observe that the nucleotide situated at position 273 is guanine (G), the amino acid resulting from the GTA codon thus being valine.

Now, the reading of the nucleotide corresponding to this position 273 in the application FR.8708090 of 10 June 1987 had led to reporting thymine (T) and leucine as amino acid resulting from the TTA codon.

Another sequence of nucleotides of the invention is characterized by its capacity of hybridization with a probe formed from the sequence (III) showing the following chain arrangement:

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534  
 535  
 536  
 537  
 538  
 539  
 540  
 541  
 542  
 543  
 544  
 545  
 546  
 547  
 548  
 549  
 550  
 551  
 552  
 553  
 554  
 555  
 556  
 557  
 558  
 559  
 560  
 561  
 562  
 563  
 564  
 565  
 566  
 567  
 568  
 569  
 570  
 571  
 572  
 573  
 574  
 575  
 576  
 577  
 578  
 579  
 580  
 581  
 582  
 583  
 584  
 585  
 586  
 587  
 588  
 589  
 590  
 591  
 592  
 593  
 594  
 595  
 596  
 597  
 598  
 599  
 600  
 601  
 602  
 603  
 604  
 605  
 606  
 607  
 608  
 609  
 610  
 611  
 612  
 613  
 614  
 615  
 616  
 617  
 618  
 619  
 620  
 621  
 622  
 623  
 624  
 625  
 626  
 627  
 628  
 629  
 630  
 631  
 632  
 633  
 634  
 635  
 636  
 637  
 638  
 639  
 640  
 641  
 642  
 643  
 644  
 645  
 646  
 647  
 648  
 649  
 650  
 651  
 652  
 653  
 654  
 655  
 656  
 657  
 658  
 659  
 660  
 661  
 662  
 663  
 664  
 665  
 666  
 667  
 668  
 669  
 670  
 671  
 672  
 673  
 674  
 675  
 676  
 677  
 678  
 679  
 680  
 681  
 682  
 683  
 684  
 685  
 686  
 687  
 688  
 689  
 690  
 691  
 692  
 693  
 694  
 695  
 696  
 697  
 698  
 699  
 700  
 701  
 702  
 703  
 704  
 705  
 706  
 707  
 708  
 709  
 710  
 711  
 712  
 713  
 714  
 715  
 716  
 717  
 718  
 719  
 720  
 721  
 722  
 723  
 724  
 725  
 726  
 727  
 728  
 729  
 730  
 731  
 732  
 733  
 734  
 735  
 736  
 737  
 738  
 739  
 740  
 741  
 742  
 743  
 744  
 745  
 746  
 747  
 748  
 749  
 750  
 751  
 752  
 753  
 754  
 755  
 756  
 757  
 758  
 759  
 760  
 761  
 762  
 763  
 764  
 765  
 766  
 767  
 768  
 769  
 770  
 771  
 772  
 773  
 774  
 775  
 776  
 777  
 778  
 779  
 780  
 781  
 782  
 783  
 784  
 785  
 786  
 787  
 788  
 789  
 790  
 791  
 792  
 793  
 794  
 795  
 796  
 797  
 798  
 799  
 800  
 801  
 802  
 803  
 804  
 805  
 806  
 807  
 808  
 809  
 810  
 811  
 812  
 813  
 814  
 815  
 816  
 817  
 818  
 819  
 820  
 821  
 822  
 823  
 824  
 825  
 826  
 827  
 828  
 829  
 830  
 831  
 832  
 833  
 834  
 835  
 836  
 837  
 838  
 839  
 840  
 841  
 842  
 843  
 844  
 845  
 846  
 847  
 848  
 849  
 850  
 851  
 852  
 853  
 854  
 855  
 856  
 857  
 858  
 859  
 860  
 861  
 862  
 863  
 864  
 865  
 866  
 867  
 868  
 869  
 870  
 871  
 872  
 873  
 874  
 875  
 876  
 877  
 878  
 879  
 880  
 881  
 882  
 883  
 884  
 885  
 886  
 887  
 888  
 889  
 890  
 891  
 892  
 893  
 894  
 895  
 896  
 897  
 898  
 899  
 900  
 901  
 902  
 903  
 904  
 905  
 906  
 907  
 908  
 909  
 910  
 911  
 912  
 913  
 914  
 915  
 916  
 917  
 918  
 919  
 920  
 921  
 922  
 923  
 924  
 925  
 926  
 927  
 928  
 929  
 930  
 931  
 932  
 933  
 934  
 935  
 936  
 937  
 938  
 939  
 940  
 941  
 942  
 943  
 944  
 945  
 946  
 947  
 948  
 949  
 950  
 951  
 952  
 953  
 954  
 955  
 956  
 957  
 958  
 959  
 960  
 961  
 962  
 963  
 964  
 965  
 966  
 967  
 968  
 969  
 970  
 971  
 972  
 973  
 974  
 975  
 976  
 977  
 978  
 979  
 980  
 981  
 982  
 983  
 984  
 985  
 986  
 987  
 988  
 989  
 990  
 991  
 992  
 993  
 994  
 995  
 996  
 997  
 998  
 999  
 1000



991  
 1081  
 1171  
 1261  
 1351  
 1441  
 1531  
 1621  
 1711  
 1801

AAT AAG ACA TAT CCA ATT CAG CCA GTT GGT CAA CTA ACA AGG GAA GTT TAT ACC GAG CCA TTA ATT TTA CCA TTA CAG TCT  
 GAG GCT CAA TTA CCI ACT TTT AAC GTT ATG GAG AGC AGC ACA ATT AGA AAT CCT CAT TTA TTT GAT ATA TTT GAT AAT CTT ACA ATC TTT  
 ACC GAT TGG TTT AGT GTT GGA CGC AAT TTT TAT TGG GGA GGA CAT CCA GTA ATA TCT AGC CTT ATA GGA GGT GGT AAC ATA ACA TCT CCT  
 ATA TAT GGA AGA GAG GCG AAC CAG GAG CCT CCA GAG TCC TTT ACT TTT AAT GGA CCG GTA TTT AGG ACT TTA TCA ATT CCT ACT TTA CCA  
 TTA TTA CAG CAA CCI TGC CAG CGC CAT TTT AAT TTA CBT GGT GAT GAA GGA GTA GAA TTT TCT ACA CTT ACA AAT AGC TTT ACS TAT  
 CCA GGA ACA GGT ACC GTT GAT TCT TTA ACT GAA TTA CCG CCT GAG GAT AAT AGT GTG CCA CCA CCA TAT AGT CAT CBT TTA TGT  
 CAT GCA ACT TTT GTT CAA AGA TCT GGA ACA CCT TTT TTA ACA ACT GAT GTA GTA TTT TCT TGG ACC CAT GGT AGT GCA ACT CTT ACA AAT  
 ACA AAT GAT CCA GAG AGA AAT AAT CAA ATA CCT TTA GTG GAA TTT AGA GTT TGG GAG GGC ACC TCT GTC AIT ACA GGA CCA GGA TTT  
 ACA GGA GAG GAT ATC CTT CCA AGA AAT ACC TTT GGT GAT TTT GTA TCT CTA CAA GTC AAT AAT TTA CCA AIT ACC CAA AGA TAC CBT  
 TTA AGA TTT CBT TAC GCT TCC AGT AGG GAT GCA CCA GAT TTA GTA TTA ACA GGA GCG GCA TCC ACA GGA GTG GGA GGC CAA GTT AGT GTA

1654  
 161 A16 CCI CII CAG AAA ACT ATE GAA AIA GGG YAG AOC TIA ACA TCI ABA ACA TII ACA IAI ACC GAI IIT AGI AAI CCI TIT ICA TIT  
 1981  
 162 GCI AAT CCA GAT ATA ATY GGG AIA AGI GAA CAA CCI CIA TIT GGI GCA GGI TCI ATY AGI AGC GGI GAA CII TAI AIA GAI AAA ATI  
 2071  
 163 ATY AII CIA GCA GAI GCA ACA TTY GAA GCA GAA TCT GAT TTA GAA ABA GCA CAA AAG GCG GIG AAT GCG CIG TIT ACT TCI TCC AAT  
 2181  
 164 AIC GGG TIA AAA ACC GAI GIT AGG GAT IAT CAI ATY GAT CAA GTA TCC AAT TTA GIG GAT TGI TTA ICA GAT GAA TIT TGI CIG GAI  
 2231  
 165 AAG LAA GAA TIG TCC GAG AAA GIT AAA CAT GCG AAG GGA CTC AGI GAT GAG GAG AAT IIA CII CAA GAI CCA AAG TIC ABA GGG AIT  
 2341  
 166 ACA CAA CCA GAC GGT GCG TGG AGA GGA AGI ACA GAT ATY ACC AIC CAA GAA GGA GAT GAC GTA TTC AAA GAG AAT TAC GTC ACA CIA  
 2431  
 167 GCI ACC GII GAI GAG TGC IAT CCA AGG IAT TIA TAI CAG AAA AIA GAT GAG TCG AAA TIA AAA GCI TAI ACC GGI IAT GAA TIA AGA  
 2521  
 168 IAT AIC GAA GAT AGI CAA GAC TIA GAA AIC IAT TIG AIC GCG TAC AAT GCA AAA CAC GAA ATA GIA AAT GIG CCA GGC AGG GGI TCC  
 2611  
 169 GCG CII ICA GCG CAA AGI CCA AIC GAA AIC GAA AAG TGI GAA AAA GCG AAT CAA TCG GCG CCA CAC CII GAA TGG AAT CCI GAI CIA GAI  
 2701  
 161 ICE TCG AG

In a distinctive manner, sequences of nucleotides of the invention coding for a polypeptide toxic specifically towards larvae of Lepidoptera of the Noctuidae family, and preferably toward S.littoralis comprise or are constituted by the chain arrangement (III) previously defined.

The chain arrangement (III), comprised in the sequence of nucleotides of the invention contains 2711 nucleotides. This fragment includes in particular the potential promoter of the gene of the  $\delta$ -endotoxin active on S.littoralis.

Sequences of nucleotides modified in relation to the chain arrangements (I) or (III) described above naturally enter into the framework of the present invention to the extent to which these modifications do not generate appreciable variations of the toxicity of the polypeptide coded by the modified sequence towards S.littoralis.

These modifications may, for example, consist of deletions, substitutions, recombinations.

Thus, the sequences of nucleotides (I) and (III) contain at their position 611 a variable nucleotide corresponding to adenine (A) in the sequence (I) and to cytosine (C) in the sequence (III). These nucleotides enter into the composition of the respective codons GAA and GCA which code respectively for the amino acids glutamic acid (GLU) and alanine (ALA) in the respective sequences II and IV.

Similarly, any sequence of nucleotides which can hybridize with that of the chain arrangements (I) or (III) such as obtained by reverse enzymatic transformation of the corresponding RNA or even by chemical synthesis also enter into the framework of the definitions of the invention.

The sequence of nucleotides of formula (III) starts with a ATG initiation codon situated at position 241 and which represents the start of an open reading frame of 2470 nucleotides.

The invention also relates to a sequence of nucleotides characterized in that it codes for a polypeptide containing the sequence (IV) of amino acids below:

271 MET GLU GLU ASN ASN GLN ASN GLN CYS ILE  
 PRO TYR ASN CYS LEU SER ASN PRO GLU GLU VAL LEU LEU ASP GLY GLU ARG ILE SER THR GLY ASN SER SER ILE ASP ILE SER LEU SER  
 301 LEU VAL GLN PHE LEU VAL SER ASN PHE VAL PRO GLY GLY GLY PHE LEU VAL GLY LEU ILE ASP PHE VAL TAP GLY ILE VAL GLY PRO SER  
 431 GLN TAP ASP ALA PHE LEU VAL GLN ILE GLU GLN LEU ILE ASN GLU ARG ILE ALA GLU PHE ALA ARG ASN ALA ALA ILE ALA ASN LEU GLU  
 541 GLY LEU GLY ASN ASN PHE ASN ILE TYR VAL GLU ALA PHE LYS GLU TAP GLU GLU ASP PRO ASN ASN PRO ALA THR ARG THR ARG VAL ILE  
 631 ASP ARG PHE ARG ILE LEU ASP GLY LEU LEU GLU ARG ASP ILE PRO SER PHE ARG ILE SER GLY PHE GLU VAL PRO LEU LEU SER VAL TYR  
 721 ALA GLN ALA ALA ASN LEU HIS LEU ALA ILE LEU ARG ASP SER VAL ILE PHE GLY GLU ARG TAP GLY LEU THR THR ILE ASN VAL ASN GLU  
 811 ASP TYR ASN ARG LEU ILE ARG HIS ILE ASP GLU TYR ALA ASP HIS CYS ALA ASN THR TYR ASN ARG GLY LEU ASN ASN LEU PRO LYS SER  
 901 THR TYR GLN ASP TAP ILE THR TYR ASN ARG LEU ARG ARG ASP LEU THR VAL LEU ASP ILE ALA ALA PHE PHE PRO ASN TYR ASP

331

ASN ARG ARG TYR PRO ILE GLN PRO VAL GLY GLN LEU THR ARG GLU VAL TYR THR ASP PRO LEU ILE ASN PHE ASN PRO GLN LEU GLN SER  
 1081 VAL ALA GLN LEU PRO THR PHE ASN VAL MET GLU SER SER ALA ILE ARG ASN PRO HIS LEU PHE ASP ILE LEU ASN ASN LEU THR ILE PHE  
 1171 THR ASP TAP PHE SER VAL GLY ARG ASN PHE TYR IAP GLY GLY HIS ARG VAL ILE SER SER LEU ILE GLY GLY ASN ILE THR SER PRO  
 1261 ILE TYR GLY ARG GLU ALA ASN GLN GLU PRO PRO ARG SER PHE THR PHE ASN GLY PRO VAL PHE ARG THR LEU SER ILE PRO THR LEU ARG  
 1331 LEU LEU GLN PRO CYS GLN ARG HIS HIS PHE ASN LEU ARG GLY GLY VAL GLU PHE SER THR PRO THR ASN SER PHE THR IYH  
 1441 ARG GLY ARG GLY THR VAL ASP SER LEU THR GLU LEU PRO PRO GLU ASP ASN SER VAL PRO PRO ARG GLU GLY TYR SER HIS ARG LEU CYS  
 1531 HIS ALA THR PHE VAL GLN ARG SER GLY THR PRO PHE LEU THR THR GLY VAL VAL PHE SER TAP THR HIS ARG SER ALA THR LEU THR ASN  
 1621 THR ILE ASP PRO GLU ARG ILE ASN GLN ILE PRO LEU VAL LYS GLY PHE ARG VAL TAP GLY GLY THR SER VAL ILE THR GLY PRO GLY PHE  
 1711 THR GLY GLY ASP ILE LEU ARG ARG ASN THR PHE GLY ASP PHE VAL SER LEU GLN VAL ASN ILE ASN SER PRO ILE THR GLN ARG TYR ARG  
 1801 LEU ARG PHE ARG TYR ALA SER SER ARG ASP ALA ARG VAL ILE VAL LEU THR GLY ALA ALA SER THR GLY VAL GLY GLN VAL SER VAL

1131 ALA MET PRO LEU GLN LYS THR MET GLU ILE GLY, GLU ASN LEU THR SER ARG THR PHE ARG TYR THR ASP PHE SER, ASN PRO PHE SER PHE  
 1381 ARG ALA ASN PRO ASP ILE ILE GLY ILE SER GLU GLN PRO LEU PHE GLY ALA GLY SER ILE SER SER GLY GLU LEU TYR ILE ASP LYS ILE  
 2071 GLU ILE ILE LEU ALA ASP ALA THR PHE GLU ALA, GLU SER ASP LEU GLU ARG ALA GLN LYS ALA VAL ASN ALA LEU PHE THR SER SER ASN  
 2161 GLN ILE GLY LEU LYS THR ASP VAL THR ASP TYR, HIS ILE ASP GLN VAL SER ASN LEU VAL ASP CYS LEU SER ASP GLU PHE CYS LEU ASP  
 2231 GLU LYS ARG GLU LEU SER GLU LYS VAL LYS HIS ALA LYS ARG LEU SER ASP, GLU ARG ASN LEU LEU GLN ASP PRO ASN PHE ARG GLY ILE  
 2341 ASN ARG GLN PRO ASP ARG GLY TRP ARG GLY SER THR ASP ILE THR ILE GLN GLY GLY ASP ASP VAL PHE LYS GLU ASN TYR VAL THR LEU  
 2431 PRO GLY THR VAL ASP GLU CYS TYR PRO THR TYR LEU TYR GLN LYS ILE ASP GLU SER LYS LEU LYS ALA TYR THR ARG TYR GLU LEU ARG  
 2521 GLY TYR ILE GLU ASP SER GLN ASP LEU GLU ILE TYR LEU ILE ALA TYR ASN ALA LYS HIS GLU ILE VAL ASN VAL PRO GLY THR GLY SER  
 2611 LEU TRP PRO LEU SER ALA GLN SER PRO ILE GLY LYS CYS GLY GLU PRO ASN ARG CYS ALA PRO HIS LEU GLU TRP ASN PRO ASP LEU ASP  
 271 CYS SER CYS

The invention also relates to recombinant expression and cloning vectors comprising more particularly at least one sequence of nucleotides such as that defined above, in particular at least a part of the sequence of about 3 kb.

5 A specific recombinant vector is, for example, a plasmid containing the HindIII-PstI fragment of the sequence of nucleotides of the invention, inserted in a vector pUC9. A first preferred vector is the plasmid pHT71, the construction of which is reported in the assemblies below, which comprises a HindIII-PstI DNA fragment according  
10 to the invention constituted uniquely of DNA derived from the strain aizawai 7-29.

Another recombinant vector is constituted by the plasmid pHT 671, the construction of which is given in figure 4. This plasmid contains a chimeric HindIII-PstI fragment, obtained by fusing a HindIII-HindII fragment of 1.1 kb derived from the strain entomocidus 6-01  
15 with a HincII-PstI fragment of 1.9 kb derived from the strain aizawai 7-29.

The modified bacterial strains which contain one of the nucleotide sequences defined above or also a recombinant expression  
20 vector and cloning previously defined, and preferably the plasmid pHT671 or the plasmid pHT71, also enter into the framework of the invention.

The invention also relates to a polypeptide toxic towards larvae of Lepidoptera and in a preferential manner towards  
25 S.littoralis, which attack cotton leaves or other crops such as those listed above, characterized in that it is capable of forming an immunological complex with antibodies directed against polypeptides with larvicidal activity towards S.littoralis.

The invention relates more particularly to the NH<sub>2</sub>-terminal  
30 part of this polypeptide which contains the larvicidal activity.

The extremity of the active NH<sub>2</sub>-terminal part corresponds to the sequence (II) of amino acids given above.

A preferred polypeptide of the invention is that which  
35 corresponds to this sequence (II) and corresponds to the sequence (IV) of amino acids given in the preceding pages. This polypeptide corresponding to the sequence (IV) contains 823 amino acids. Its calculated molecular mass is 92906 Da.

This sequence of  $\delta$ -endotoxin was compared with amino acid sequences of  $\delta$ -endotoxins derived from other strains of B.thuringiensis active on the Lepidoptera and the genes of which have been isolated and sequenced previously : the  $\delta$ -endotoxins in question are those of the strains kurstaki HD1 (19), kurstaki HD73 (20), berliner 1715 (21) and (22) Sotto (23) and aizawai IPL7 (24).

The results of these comparisons indicate that all are different in the second quarter of the molecule (amino acids 281 to 620) whereas the  $\text{NH}_2$ -terminal part (amino acids 1 to 280) and the  $\text{COOH}$ -terminal domain (amino acids 621 to 1175) of the protein are highly conserved and differ only by several amino acids. On the other hand, the  $\delta$ -endotoxin corresponding to the sequence (IV) above shows considerable differences from the other  $\delta$ -endotoxins both in the  $\text{NH}_2$ -terminal part (amino acids 1 to 280) and in the second quarter of the molecule (amino acids 281 to 620). The results of these comparisons indicate again that the  $\text{NH}_2$ -terminal half of the molecule (amino acids 1 to 620) which corresponds to the toxic fraction of the protein only show 46% homology with the other  $\delta$ -endotoxins. The most important differences are located in the second half of the toxic part of the molecule (amino acids 281 to 620) with only 36% of identical amino acids, the  $\text{NH}_2$ -terminal part (amino acids 1 to 280) itself showing 58% of amino acids identical with the other  $\delta$ -endotoxins. Such considerable differences have never been observed up to now in the  $\text{NH}_2$ -terminal part of the toxic fraction of the molecule among the  $\delta$ -endotoxins active on the Lepidoptera.

In order to obtain a sequence of nucleotides entering into the framework of the invention, coding for at least the active part of a polypeptide showing a specific toxicity towards larvae of Lepidoptera of the Noctuidae family, and preferably towards S.littoralis, recourse is had, in conformity with the invention, to the following steps, namely:

- the carrying out of a molecular hybridization between, on the one hand, a nucleotide sequence of a strain of B.thuringiensis active against S.littoralis and, on the other, at least two nucleotide sequences, used as probes, derived from the 5' part of a restriction



fragment of a gene for  $\delta$ -endotoxin of B.thuringiensis, this part coding for the NH<sub>2</sub>-terminal part of the polypeptide active on the larvae of Lepidoptera, and from the 3' part of this fragment coding for the COOH part of the polypeptide,

- 5 - the isolation of the hybrid fragment,
- its cloning in a vector, followed by its purification.

In an advantageous manner, the hybridization probes utilized are obtained from a gene for the  $\delta$ -endotoxin derived from the strain aizawai 7-29 coding for a protein of 130 kDa, active against P.brassicae and inactive towards S.littoralis, this gene having been  
10 cloned in the recombinant plasmid pHTA2.

In an embodiment of the preceding procedure, the fragment recombined with the vector in the cloning step is elaborated from a HindIII-PstI restriction fragment derived from a single strain of B.thuringiensis, preferably aizawai 7-29. In particular, this fragment  
15 is carried preferentially by the recombinant plasmid pHTA6 such as isolated with the aid of a probe constituted by a PvuII fragment of 2 kb of the plasmid pBT15-88 corresponding to the internal part of a gene for the chromosomal crystal of the strain berliner 1715,  
20 starting from transforming clones containing nucleotide sequences derived from B.thuringiensis strains active against larvae of Lepidoptera, inter-alia of S.littoralis.

In another embodiment of the invention, the fragment recombined with the vector in the cloning step is elaborated from  
25 several sequences of nucleotides derived from recombinant vectors containing sequences of nucleotides from at least two different strains of B.thuringiensis, possessing the same restriction maps and themselves containing all or part of the sequences of nucleotides capable of coding for a polypeptide active, in a preferential manner, against  
30 S.littoralis.

In this case, the recombined fragment used in the cloning step is a fragment of about 3 kb, advantageously elaborated from a HindIII-HincII restriction fragment of about 1.1 kb derived from the entomocidus 6-01 strain and a HincII-PstI fragment of about 1.9 kb  
35 from the aizawai 7-29 strain. It corresponds to a truncated gene for  $\delta$ -endotoxin.

The HindIII-HincII and HincII-PstI restriction fragments are carried more especially by the respective recombinant plasmids pHTE6 and pHTA6 such as isolated with the aid of the probe constituted by the PvuII fragment mentioned above.

5           The study of the toxicity towards the larvae of Lepidoptera of the bacterial strains modified with the aid of the sequences of nucleotides defined above, has made it possible to demonstrate their high toxic activity, in particular with regard to the caterpillars of S.littoralis.

10           This activity was estimated from the point of view of the specificity index corresponding to the ratio

LC50 S.littoralis

---

LC50 P.brassicae

15           in which "LC50" represents the lethal concentration killing 50% of the larvae in 72 hours.

          The utilization of such an index makes it possible to evaluate the activity of the bacterial strains studied without having to consider the level of expression of the polypeptides.

20           The results obtained, which are reported in the examples which follow, and the values of LD50 which are deduced from them, have shown that the bacterial strains modified according to the invention show a more specific toxic activity towards S.littoralis than the native crystal proteins of the strains aizawai 7-29 or  
25   berliner 1715.

          Therefore, the invention relates to the use of these modified strains, of recombinant vectors containing the nucleotide sequences defined above, in particular the plasmid pHT671 and the plasmid pHT71, and these sequences themselves for the elaboration of larvicidal  
30   compositions preferentially toxic towards S.littoralis.

          The larvicidal compositions of the invention are thus characterized in that they contain an efficaceous quantity of polypeptides such as defined above or expressed by the nucleotide sequences mentioned above.

35           In order to produce these polypeptides the truncated genes

for  $\delta$ -endotoxin corresponding to the nucleotide sequences of the invention are advantageously implemented.

These genes can be used to produce in excess the toxic polypeptide in microorganisms permitting the expression of the above recombinant vectors. Suitable strains of microorganisms include E.coli or also B.subtilis.

These truncated genes may be reintroduced into the strains of B.thuringiensis or into related species such as B.cereus, according to the standard techniques, for example, by transformation, conjugation or transduction. These techniques make it possible to produce the toxic polypeptide in large quantity without first having to modify the natural region of the promoter for the  $\delta$ -endotoxin genes of B.thuringiensis.

This transformation may be carried out by using methods derived from the transformation of the protoplasts of B.subtilis according to (11) or of the vegetative cells of B.thuringiensis as described in (12).

The introduction of recombinant plasmids by a system of the conjugation type may be carried out by using B.thuringiensis as host strain and B.subtilis or Streptococcus faecalis as strains of the donor type by operating according to (13) and (14).

As a variant, the sequences of nucleotides are introduced into microorganisms living in the environment or in association with the plants and capable of expressing recombinant vectors containing these sequences. The introduction may be carried out in microorganisms such as Pseudomonas by working according to the procedure described in (17), by the intermediary of plasmid vectors containing the transposon Tn5 and the gene for the toxin, or Azospirillum or Rhizobium by means of the intermediary of suicide vectors derived from the plasmid RP4 and of mobilizing plasmids functional in Gram negative bacteria (for example, pRK2013) according to the procedures described in (18).

The truncated genes are alone in the strains of Bacilli or, as a variant, are associated with different  $\delta$ -endotoxin genes which makes it possible to obtain crystals synthesized by these strains

specifically toxic towards given species of Noctuidae, or toxic both towards the Noctuidae and insects sensitive to other  $\delta$ -endotoxins.

These recombinations, carried out in vitro or in vivo with the nucleotide sequences of the invention and other  $\delta$ -endotoxin genes showing different toxic specificities, lead to the construction of new genes coding for novel hybrid toxic proteins exhibiting a large spectrum of activity towards insects. These new genes and these novel proteins also enter into the framework of the invention.

In these applications, the nucleotide sequences of the invention may be transferred and expressed in plants sensitive to S.littoralis in order to diminish the devastation caused by these insects.

Among the plants to be protected, mention should be made of : cotton, clover, the tomatoe and alfalfa.

The transfer of the truncated gene into cotton plants may be carried out by transformation involving strains such as Agrobacterium as described in (15).

In addition, the invention relates to the plant cells, the plants and the seeds containing the nucleotide sequences defined above.

The plant cells according to the invention are cells, the genome of which after transformation by a non-essentially biological procedure possesses in a stable manner a sequence of nucleotides capable of expressing a polypeptide toxic towards S.littoralis, such as that defined above. The invention also relates to the plant cells derived from their division.

The plants according to the invention are plants transformed by a non-essentially biological procedure, having in particular as predator S.littoralis, the genome of which possesses in a stable manner a sequence of nucleotides such as that defined above, capable of expressing a polypeptide toxic towards S.littoralis. The plants in question are also plants derived from their reproduction, their multiplication or hybrid crosses.

In accordance with another feature, the invention relates to plants having in particular as predator S.littoralis, possessing

in addition to their initial phenotypic and genotypic characters the property of expressing a polypeptide toxic preferentially towards S.littoralis, this property resulting from the insertion in their genome by means of genetic manipulation of a sequence of nucleotides capable of expressing the said polypeptide.

In addition, the invention relates to seeds capable of giving rise to a plant such as that defined above or derived from such a plant, characterized in that they have integrated into their genome by genetic manipulation a nucleotide sequence described above.

Other characteristics and advantages of the invention will become apparent in the course of the description and in referring to the examples in which :

- figure 1 presents the restriction map of the plasmids pHTA6 and pHTE6,
- figure 2, the restriction map of a gene for a crystal protein of the aizawai 7-29 strain cloned in the plasmid pHTA2 and defining the DNA fragments which are used as probe,
- figure 3 shows the fragment of 6.6 kb cloned in pHTA6 and the result of a hybridization carried out between this fragment and the probes described in figure 2,
- figure 4, the restriction map of the plasmid pHT671, and
- figure 5, the photographs of the immunodiffusion tests.

The hybridization experiments carried out for the implementation of the invention were performed at 42°C for 24 h in a solution containing 5 x SSC, 30% formamide and 1 Denhardt (7) in the presence of the DNA probe labelled with <sup>32</sup>P. The filters are washed at 42°C, 20 mn, by using successively the following solutions : 5 x SSC in 30% formamide, 5 x SSC, 2 x SSC, 1 x SSC and 0.5 x SSC before drying at room temperature.

EXAMPLE 1 - Construction of a DNA sequence of about 3 kb containing a hybrid gene of an insecticidal toxin.

This construction comprises :

- 1/ the preparation of gene banks of B.thuringiensis
- 2/ the selection and characterization of transforming clones containing the genes of a crystal protein and nucleotide sequences responsible

for the larvicidal activity,

3/ in vitro recombination of these sequences in a cloning vector with construction of the plasmid pHT671.

These different steps are carried out as follows :

5 1/ Preparation of gene banks of B.thuringiensis.

The total DNA of the aizawai 7-29 and entomocidus 6-01 strains of Bacillus thuringiensis is purified by using the method reported in (1) and 50 µg of each purified DNA are completely digested with the restriction enzyme PstI.

10 The DNA digested by PstI is analysed by horizontal electrophoresis on a 0.8% agarose gel and DNA fragments of a size of 5 to 8 kb are recovered from the agarose gels by electroelution in a manner described in (2).

15 The purified DNA fragments of 5-8 kb of the aizawai 7-29 strain are ligated to the DNA of the cloning vector pUC18 digested by PstI according to (3).

20 The purified DNA fragments of 5-8 kb of the entomocidus 6-01 chain are ligated to the DNA of the cloning vector pUC9 digested by PstI. The cells of E.coli JM83 are transformed with the ligation mixture as described in (4).

The transforming clones of E.coli are selected on LB medium containing 100 µg/ml of ampicillin.

2/ Isolation and characterization of the transforming clones containing the genes for a crystal protein.

25 A/ Screening of the transformed E.coli cells with the aid of an internal fragment of a gene of the crystal protein labelled with <sup>32</sup>P, used as probe :

30 Transforming clones containing recombinant plasmids carrying the gene for the crystal are detected by colony hybridization according to the method described in (5), by using as probe a PvuII fragment of 2 kb of the pBT 15-88 plasmid corresponding to an internal part of the gene for the crystal protein located on the chromosome of the berliner 1715 strain.

35 B/ Characterization of the recombinant plasmids present in the clones which react with the above probe.

Two recombinant plasmids, pHTA6 and pHTE6, isolated respectively from gene banks constructed from the strains aizawai 7-29 and entomocidus 6-01, show a homology with this probe. In each case, a DNA fragment of about 6.6 kb was cloned.

5 The restriction map of the two plasmids is given in figure 1. The comparison of the restriction sites shows that the two DNA fragments cloned appear to be identical.

10 In order to delimit the sequences corresponding to the gene for the  $\delta$ -endotoxin, different DNA fragments labelled with  $^{32}\text{P}$ , derived from a gene of the crystal previously characterized, and cloned in the recombinant plasmid pHTA2, are utilized as probes. This latter gene for the crystal also derived from the aizawai 7-29 strain codes for a protein of 130 kd active against P.brassicae but not against S.littoralis. This type of gene possesses the same restriction map  
15 as the gene for the  $\delta$ -endotoxin derived from the berliner 1715 strain. In figure 2 is shown the restriction map of this gene for the crystal protein of the aizawai 7-29 strain cloned in the plasmid pHTA2. The thick lines shown above the map correspond to the fragments used as hybridization probes.

20 The plasmids pHTA6 and pHTE6 are hydrolysed by different restriction endonucleases, analysed by horizontal electrophoresis on a 0.8% agarose gel and hybridized with the different probes.

The transfer of the DNA to nitrocellulose filters is carried out according to the method of SOUTHERN described in (6). The  
25 hybridization is conducted at 42°C for 24 hours in a solution containing : 5 x SSC, 30% formamide and a 1x Denhardt mixture described in (7) in the presence of a DNA probe labelled with  $^{32}\text{P}$ . The filters are then washed at 42°C for 20 minutes, by using successively the following solutions : 5 SSC in 50% formamide, 5 SSC, 2 SSC, 1 SSC  
30 and 0.5 SSC before being dried at room temperature.

The results of these hybridization experiments are summarized in figure 3. It appears that each extremity of the cloned DNA fragments of 6.6 kb shows a homology with the probes. The PstI-KpnI fragment of 1.5 kb reacting with the probe No. 3 corresponds to the 3' end  
35 of a gene of the crystal protein present in both the aizawai 7-29

and entomocidus 6-01 strains. As pointed out in figure 3, the probes No. 1 and No. 2 corresponding to the 5' end of the gene for the  $\delta$ -endotoxin of pHTA2 hybridize with the HindIII-HincII fragment of 1.1 kb contained in the plasmid pHTA6. The probe No. 3 which covers the 3' end of the gene of the  $\delta$ -endotoxin of pHTA2 hybridizes with the HindIII-PstI fragment of 0.4 kb contained in the plasmid pHTA6. It should be noted that a weak hybridization signal is obtained with the probe No. 2 whereas the two other probes give easily detectable signals.

From these results, the inventors have established that the HindIII-PstI DNA fragment of 3 kb corresponds to a large part of a gene for the  $\delta$ -endotoxin which commences close to the central HindIII site. It seems clear in the light of results of the hybridization experiments that the gene for the  $\delta$ -endotoxin shows substantial differences from those characterized in the prior art. On the basis of these results it was decided to clone the HindIII-PstI fragment of 3 kb in the vector pUC9.

3/ Construction of the plasmid pHT 671 containing a hybrid gene of the reconstituted insecticidal toxin.

The HindIII-HincII DNA fragment of 1.1 kb derived from the plasmid pHTE6 and the HincII-PstI DNA fragment of 1.9 kb derived from the plasmid pHTA6 are purified on agarose gels.

Equal amounts of the two purified DNA fragments and the DNA of pUC9 digested with HindIII and PstI are mixed and ligated. The ligation mixture is used to transform competent cells of E.coli JM83, then the transformed E.coli cells are selected on LB medium containing ampicillin. One of the interesting recombinant clones examined contains a plasmid designated by pHT671, the restriction map of which was determined and is shown in figure 4. This plasmid (pHT671) contains a DNA fragment of 3 kb inserted in the vector pUC9. This DNA sequence has the same restriction map as the HindIII-PstI fragments of 3 kb contained in the plasmids pHTA6 and pHTE6, but corresponds to a reconstituted DNA molecule constructed by in vitro recombination from DNA sequences derived from the aizawai 7-29 strains on the one hand and entomocidus 6-01 on the other.



EXAMPLE II : Study of the nucleotide sequence of the promoter region and of the region coding for the NH<sub>2</sub>-terminal part of the  $\delta$ -endotoxin active against the Noctuidae.

5 The HindIII-HincII fragment of pHT671 is sequenced in conformity with the method described in (8) by using a M13 system. In order to obtain partially overlapping cloned DNA fragments which will be used in the sequencing of the DNA, recourse is had to the method of subcloning by deletion in M13, developed by DALE et al (9).

10 The sequence of 940 nucleotides of the HindIII-HincII fragment which has a length of about 1 kilobase corresponds to the chain arrangement I above.

15 The analysis of this sequence shows that the largest open reading frame starts at position 241 and that a potential site of binding to the ribosomes, GGAGG, is found six base pairs upstream from this ATG codon (position 230 to 235). The region localized between the nucleotides 137 and 177 (position -103 to -63 upstream from the ATG codon) is strongly homologous with the region present upstream from the gene for the crystal of the strain kurstaki HD1 Dipel (BTK) sequenced by WONG et al (1983) and described in (16) and the authors of which have shown that it contains three promoters BtI, BtII, and Ec, functional in B.thuringiensis and E.coli, respectively. The comparison between the amino acid sequences deduced from the first 750 nucleotides of the genes of BTK and pHT671, show that these polypeptides exhibit significant differences at the level of the N-terminal half of the active part derived from the protoxin (only 66% strict homology). It is important to note that it is the first time that a gene for the  $\delta$ -endotoxin isolated from a strain active against the Lepidoptera codes for a polypeptide which shows substantial differences in this region. In fact, this N-terminal domain appears to be strongly conserved (more than 97% of strict homology) among all of the genes for the crystal active on Lepidoptera which have been sequenced hitherto. Moreover, the inventors have shown that the degree of variability is of the same order if the nucleotide sequences of pHT671 and other genes of the Lepidoptera type are considered.

EXAMPLE III : Construction of a DNA sequence of about 2.7 kb containing a gene for a larvicidal toxin.

In order to achieve this construction the DNA of the aizawai 7-29 strain of B.thuringiensis was used up to the step for the production of the plasmid pHTA6 as described in Example I.

The HindIII-PstI fragment of about 2.7 kb obtained from the plasmid pHTA6 was then subcloned in the vector pUC9, previously hydrolysed by the restriction enzymes HindIII-PstI in order to give the plasmid pHT71.

EXAMPLE IV : Study of the sequence of nucleotides constituting the plasmid pHT71 coding for a polypeptide toxic towards the larvae of Lepidoptera of the family of the Noctuidae.

The HindIII-PstI fragment of 2.7 kb of pHTA6, which was subcloned in pHT71, was sequenced by means of the technique of Sanger et al. (8) using the phage M13mpl9 and the subcloning system by deletions developed by Dale et al (9). This system makes it possible to obtain M13 phages containing a series of partially overlapping DNA fragments which can be utilized for sequencing the DNA.

The sequence of nucleotides of this 2.7 kb fragment which corresponds to the chain arrangement (III) given above, was determined on the 2 DNA strands, with the exception of the last 212 nucleotides (position 2500 to 2711) which were sequenced only on a single strand.

The nucleotide sequence of this HindIII-PstI fragment has a length of 2711 nucleotides. This fragment contains the potential promoter as well as the largest part of the gene for the  $\delta$ -endotoxin active on S.littoralis.

EXAMPLE V : Study of the specific toxicity of the recombinant clones of E. coli JM83 (pHT671) and JM83 (pHT71) against S.littoralis.

The toxicity of the recombinant clones of E.coli JM83 containing pHT671 and of E.coli JM83 containing pHT71 was determined by biological tests on caterpillars of the P.brassicae and S.littoralis species as described by LECADET and MARTOURET in (10). The results were compared with the specific toxicity of the native crystal proteins purified from the strains berliner 1715 and aizawai 7-29, entomocidus 6-01 B.cereus 569 (containing the plasmid pBT45, pAMB1) against the

two species of insects. The specific toxicity of the recombinant clone and of the strains of B.thuringiensis is expressed in terms of "specificity index" previously defined.

The results obtained are reported in table 1 below.

5 In this table, for E.coli strains, the concentration 1 corresponds to a 14 hours bacterial culture concentrated 20 times, disintegrated by ultrasound ; for the B.thuringiensis strains the concentration is expressed in  $\mu\text{g}$  of crystal protein per  $\mu\text{l}$  of preparation. The toxic activity of the preparations was tested by  
10 the forced ingestion with 5  $\mu\text{l}$  of preparation on caterpillars at the fifth stage of development, or by a technique of free ingestion utilizing larvae at the second stage of development.

15

20

25

30

35

TABLE 1

Comparative toxicity of a recombinant clone and two strains of B. thuringiensis towards S. littoralis and P. brassicae.

5	Strains and plasmids	S.littoralis		P.brassicae		Specificity index LC50 <u>S.littoralis</u> LC50 <u>P.brassica</u>
		LC50 2nd larval stage	LC50 5th larval stage	LC50 5th larval stage	LC50 5th larval stage	
10	JM83 (pUC18)	> 1	> 1	> 1		-
	JM83 (pHT671)	0,04	0,13	0,72		0,2
	JM83 (pHTA2)	> 1	> 1	0,03		> 30
15	JM83 (pHTA4)	> 1	> 1	> 1		-
	JM83 (pHT71)	ND	0,5	> 1		< 0,5
	berliner 1715 native crystals	ND	0,11	0,007		15,7
20	aizawai 7.29 native crystals	ND	0,02	0,04		0,5
	entomocidus 601 native crystals	ND	0,028	0,012		2,3
25	B.cereus 569 (pBT45.pAM $\beta$ 1)	ND	0,38	0,054		7

30

35

Examination of the LC50 values summarized in this table 1 shows that the protein extracts of the recombinant clones JM83 (pHT671) and JM83 (pHT71) are preferentially toxic against S.littoralis.

Secondly, a comparison of the values of the specificity index shows that the larvicidal activity of the recombinant clones is more specific by a factor of 2.5 times towards S.littoralis than the native crystal proteins of the aizawai strain. Moreover, the recombinant clones of JM83 (pHT671) and JM83 (pHT71) are very active against another insect of the family of the Noctuidae, Mamestra brassicae (in the case of the clone JM83 (pHT671) for example, the LC50 value is 0.02, utilizing larvae at the second stage of development).

These two results show that the gene for the larvicidal toxin constructed and cloned in the plasmids pHT671 and pHT71 codes for a protein specifically active against S.littoralis.

Other preparations obtained from recombinant clones containing plasmids carrying genes coding for other types of  $\delta$ -endotoxins (pHTA2 and pHTA4) are not active on S.littoralis : it may be seen that the plasmid pHTA2 codes for a  $\delta$ -endotoxin specifically active on P.brassicae whereas the plasmid pHTA4 codes for a  $\delta$ -endotoxin, the insect target for which has not yet been identified. It can also be seen that the crystalline inclusions produced by a strain of Bacillus cereus which has received the plasmid pBT45, one of the plasmids of the aizawai 7-29 strain which also carries a  $\delta$ -endotoxin gene (the gene of plasmid origin of the aizawai 7-29 strain), are also specifically active on P.brassicae.

Similar results are obtained by using, in the place of crude bacterial extracts, soluble protein extracts prepared from different recombinant clones of E.coli.

On the basis of the LC50 values reported in the table above and a mean individual weight of 41 mg per L5 larva (fifth larval stage) of S.littoralis, the value of the LD50 was estimated at 2.4  $\mu$ g/gram of larva for the native crystals of the aizawai 7-29 strain.

On these same bases and on the basis of equivalence factors making it possible to pass from the total bacterial mass to the

quantity of specific proteins (about 2% of the total proteins in E.coli JM83 (pHT671), the LD50 corresponding to the toxin produced by the expression in E.coli JM83 of the gene according to the invention cloned in the plasmid pHT671, was determined and estimated at a value close to 5.5 to 6 µg/gram of larva.

On these same bases and after determination of the LC50 of soluble protein extracts prepared from ground cultures of E.coli JM83 (pHT671), the value of the LD50 corresponding to the toxin present in these extracts was estimated at 4.15 µg/gram of larva.

In the two cases and particularly in the case of the ground preparations of E.coli, the calculated values of LD50 are approximate and higher than that of the native crystals, because it is not a question of a purified toxin. However, these data indicate without ambiguity that the gene expressed by pHT671 specifies a  $\delta$ -endotoxin exhibiting the specificity towards S.littoralis. In fact, the same type of estimation made with extract of E.coli JM83 (pHTA2) carrying a  $\delta$ -endotoxin gene of different specificity leads to values 30 to 50 times higher than the LD50 of the soluble extracts towards S.littoralis (135 to 350 µg/gram of larva).

The foregoing data will easily make it possible for the person skilled in the art to develop active larvicidal compositions with the proteins of the invention.

Other toxicity experiments were carried out utilizing larvae of M.brassicae, S.fruqiperda and S.littoralis at the second larval stage. The results obtained, expressed in terms of LC50 as defined for table 1, are given in table 2.

TABLE 2

ACTIVITY OF THE RECOMBINANT CLONES AGAINST  
THE LARVAE OF INSECTS OF THE FAMILY OF THE  
NOCTUIDAE: M. BRASSICAE, S. FRUGIPERDA, and  
S. LITTORALIS.

STRAINS AND PLASMIDS	INSECT LARVAE AND STAGE	<u>M. BRASSICAE</u>		<u>S. FRUGIPERDA</u>		<u>S. LITTORALIS</u>	
		LC50	2nd STAGE	LC50	2nd STAGE	LC50	2nd STAGE
JM 83 (pUC18)		NT		NT		NT	
JM 83 (pHTA2)		> 1		0,51		0,9	
JM 83 (pHT671)		0,02		0,5		0,03	
JM 83 (pHT71)		ND		ND		0,03	
JM 83 (pHTA4)		> 1		0,54		> 1	

It emerges from the examination of table 2 that the crude bacterial extracts of the recombinant clone JM83 (pHT671) are toxic towards M.brassicae and S.littoralis (the values of LC50 are 0.02 and 0.03, respectively) and weakly toxic towards S.fruqiperda (LC50 of 0.5).

The extracts of the recombinant clone E.coli JM83 (pHTA2) are weakly active towards S.fruqiperda and S.littoralis and not at all toxic towards M.brassicae. The extracts of the recombinant clone JM83 (pHTA4) are not toxic towards M.brassicae and S.littoralis and are weakly toxic toward S.fruqiperda.

These results confirm the high specific toxicity of the proteins obtained from pHT71 and pHT671 towards S.littoralis and show that this class of crystal protein is also very active towards M.brassicae.

EXAMPLE VI : Study of the specificity of the polypeptides expressed by the clones formed by introduction of the plasmids pHT671 and pHT71 into E.coli.

This study was carried out owing to immuno-diffusion tests. The results are reported in figure 5 (which includes figures 5A and 5B).

The implementation of the immuno-diffusion experiment was done in conformity with the following protocol :

Soluble extracts of proteins of E.coli clones containing the plasmids pHT671, pHTA4, pHTA2 or pHT71, pUC18 were placed in the wells Nos. 2, 3, 4, 5, 6, respectively. A sample of a solubilized purified crystal of aizawai 7-29 was placed in the well No. 1 in order to serve as positive control.

In the test recorded in figure 5A an antiserum against all of the  $\delta$ -endotoxins of aizawai 7-29, containing rabbit antibodies directed against the solubilized crystal proteins, was used and placed in the central well.

An immunoprecipitation line was observed in all of the cases except in the case of the extract of E.coli containing only the plasmid vector (well No. 6).



It was observed that the immuno-precipitation lines derived from the wells No. 4 and No. 5 cross, which shows that the products encoded by the plasmids pHTA2 and pHT71, respectively, display different antigenic determinants.

5 In the test recorded in figure 5B, the anti-serum used contained rabbit polyclonal antibodies against the crystal proteins of berliner 1715.

10 An immunoprecipitation line was observed with the extracts of E.coli JM83 (pHTA4) (well No. 3) JM83 (pHTA2) (well No. 4). On the other hand, the E.coli clones JM83 (pHT71) (well No. 5), JM83 (pHT671) (well No. 2) or JM83 (pUC9) (well No. 6) did not give immuno-precipitation.

15 It may be deduced from that that the genes for the crystal isolated in pHTA4 and pHTA2 express polypeptides having antigenic determinants in common with the proteins of the crystal of berliner 1715, a strain which is not specifically active towards S.littoralis.

20 On the other hand, the crude extracts of E.coli containing the plasmids pHT671 and pHT71 contain polypeptides having antigenic determinants in common with the crystal proteins of the aizawai 7-29 strain, which are not related immunogenically with the crystal proteins of the berliner 1715 strain.

These results confirm those given previously with respect to the specificity of the genes isolated in the plasmids pHT71 and pHT671.

25 Antigen-antibody precipitation assays have made it possible to determine the level of expression of the  $\delta$ -endotoxin genes in different recombinant clones.

30 The results obtained have shown that the crystal protein represents between 7 and 10% of the total cellular proteins of E.coli JM83 (pHTA2), between 2 and 3% in E.coli JM83 (pHT671) and between 0.5 and 1% in E.coli JM83 (pHTA4) and E.coli JM83 (pHT71).

The literature references cited in the examples  
are the following :

- 5 (1) KLIER, A.F., LECADET, M-M. and DEDONDER, R., 1973, Sequential modifications of RNA polymerase during sporogenesis in Bacillus thuringiensis, Eur. J. Biochem., 36 : 317-327.
- (2) MANIATIS, T., FRITSCH, E.F., SAMBROOK, J., 1982, Molecular cloning : A laboratory manual. Cold Spring Harbor Laboratory Press, New-York
- 10 (3) VIEIRA, J. and MESSING, J., 1982, The pUC plasmids, and M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers, Gene, 19 : 259-268.
- 15 (4) LEDERBERG, E.M. and COHEN, S.N., 1974, Transformation of Salmonella thyphimurium by plasmid deoxyribonucleic acid, J. Bacteriol., 119 : 1072-1074.
- (5) GRUNSTEIN, M. and HOGNESS, D.S., 1975, Colony hybridization, a method for the isolation of cloned DNAs that contain a specific gene, Proc. Natl. Acad. Sci. U.S.A., 72 : 3961-3965.
- 20 (6) SOUTHERN, E.M., 1975, Detection of specific sequence among DNA fragments separated by gel electrophoresis, J. Molec. Biol., 98, 503-517.
- 25 (7) DENHARDT, D.T. 1976, A membrane filter taking for the detection of complementary DNA. Biochem. Biophys. Res. Comm., 23 : 641-646.
- (8) SANGER, F., NICKLENS, S. and COULSON, A.R., 1977, DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A., 74 : 5463-5467.
- 30 (9) DALE et al. (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA, Plasmid 1985, 13 : 31-40
- 35 (10) LECADET, M.M. et MARTOURET D. 1987, Host specificity of the Bacillus thuringiensis  $\delta$ -endotoxin toward

- Lepidopteran species : Spodoptera littoralis Bdv and Pieris brassicae L, J. of Invert. Pathol., 49 (n° 1) : 37-48.
- (11)CHANG et al., 1979, High frequency transformation of Bacillus subtilis protoplasts by plasmid DNA-  
5 Mol Gen Genet 168:111 115
- (12)HEIERSON et al., 1987, Transformation of vegetative cells of Bacillus thuringiensis by plasmid DNA, Journal of Bacteriology, Mar.1987, p.1147-1152,
- 10 (13)KLIER et al., 1983, Mating between Bacillus subtilis and Bacillus thuringiensis and transfer of cloned crystal genes, Mol Gen Genet (1983) 191:257 262
- (14)LERECLUS et al., 1983, Isolation of a DNA, sequence related to several plasmids from Bacillus thuringiensis after a mating involving the Streptococcus faecalis  
15 plasmid pAM $\beta$ 1, Mol Gen Genet (1983) 191:307-313
- (15)UMBECK et al., 1987, Genetically transformed cotton (Gossypium hirsutum L.) plants - Biotechnology vol.5 March 1987.
- 20 (16)WONG et al., 1983, transcriptional and translational start sites for the Bacillus thuringiensis crystal protein gene. J. of Biol. Chem., 258 : 1960-1967.
- (17)OBUKOWICZ M.et al (1986). Tn<sup>5</sup> mediated integration of the  $\delta$ -endotoxin gene from B. thuringiensis into the  
25 chromosome of root colonizing Pseudomonas. J. Bacteriol., 168, 982-989.
- (18)SIMON, R. et al, (1983). A broad host range mobilization system for in vivo genetic engineering : transposon mutagenesis in Gram-negative bacteria. Biotechnology, 1, pp. 784-791.  
30
- (19)Schnepf et al,(1985)... The amino acid sequence of a crystal protein from Bacillus thuringiensis deduced from the DNA base sequence. J BIOL Chem 260 : 6264-6372.
- (20)Adang et al,(1985) Characterized full-length and  
35 truncated plasmid clones of the crystal protein of

Bacillus thuringiensis subsp. kurstaki HD-73 and their toxicity to Manduca sexta. Gene 36 : 289-300.

(21)Wabiko et al,(1986) Bacillus thuringiensis entomocidal protoxin gene sequence and gene product analysis. DNA 5 : 305-314.

(22)Hofte et al,(1986) Structural and functional analysis of a cloned  $\delta$ -endotoxin gene of Bacillus thuringiensis berliner 1715. Eur J Biochem 161 : 273-280.

(23)Shibano et al,(1986) Complete structure of an insecticidal crystal protein gene from Bacillus thuringiensis. In : Bacillus molecular genetics and biotechnology applications. J.Ganesan, A.T., Hoch, J.A.(eds). Academic Press 307-320.

(24)Oeda et al,(1987) Nucleotide sequence of the insecticidal protein gene of Bacillus thuringiensis strain aizawai IPL7 and its high-level expression in Escherichia coli. Gene 53 : 113-119.

20

25

30

35